DOCUMENT RESUME

ED 081 799 TM 003 160

AUTHOR Roberts, Tim
TITLE Some Utilities of Criterion Testing.
PUB DATE Jun 73
NOTE 7p.; Paper presented at Learning Disabilities Conference (Longmont, Colorado, November-December 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Criterion Referenced Tests; *Evaluation Techniques; *Norm Referenced Tests; Speeches; *Student Evaluation; *Teacher Developed Materials

ABSTRACT

Criterion-referenced testing is designed to determine the student's ability to perform a specified skill within a comprehensive task. To measure the student's degree of development or growth, a teacher-made test is used. This test is used as criterion-referenced in the measuring of a student's degree of achievement of specified behaviors within a delimited parameter. The test becomes normative based when the teacher assigns a numerical quantity to each paper for the purpose of grade distribution. With criterion-referenced tests, measuring growth can be accomplished through either positive or negative behavioral observations. With established behavioral objectives, teacher-made tests and classwork are excellent for criterion referencing. For the skilled psychometrician, normative based testing is a limited sampling of behavior that is done according to specific prerequisites. Norm testing focuses upon sorting and selecting students for placement into the educational curricula. By doing a detailed analysis of errors made on normative testing, the examiner can obtain a more detailed picture of the child than by using the actual standard score only. Thus, the skilled examiner is better able to relate testing observation to behavior displayed within the classroom setting, as well as viewing each test or subtest as to its component parts of incoming stimulus, understanding stimuli or expressive reaction. The skilled teacher must also, through observation, view a given task in the form of task analysis. From this, both the examiner and the teacher can determine the instructional level for the child and agree as to the remedial approach to be used. (DB)

ED 081799

TM 003 160

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"SOME UTILITIES OF CRITERION TESTING"

by MR. TIM ROBERTS

Paper Read at Learning
Disabilities Conference
November - December, 1972
Longmont, Colorado

Sponsored by

Northern Colorado Educational
Board of Cooperative Services
830 S. Lincoln Street
Longmont, Colorado 80501

June 1973

FILMED FROM BEST AVAILABLE COPY

"SOME UTILITIES OF CRITERION TESTING"
by Tim Roberts

The current trend in education is dealing with the individualized curriculum. This curriculum has been used quite successfully in the field of special education and now has grown into the regular classroom. With the impact of individualized instruction being felt throughout the nation, special assessment perplexities have also developed. To adequately assess the individualized curriculum, criterion-referenced testing has been introduced.

Criterion-referenced testing is new terminology for old concepts. In Ancient China, leaders were chosen by their ability to perform at specific tasks that were deemed necessary for leadership. Today with the advent of criterion-referenced testing, children are assessed in relationship to those skills that are deemed as necessary to perform a given task or to attain success in the educational process. Within the last 70 years testing has moved from criterion to the normative based test with the pendulum now beginning to swing back. Binet originally observed people's behavior and concluded that to experience success, specific objectives, skill levels or criterions must be reached before the next level of development can be attained. The observation of large groups developed a norming base upon which an individual could be compared. This comparison led to the concepts of mental age, grade equivalents and intelligence quotient.

The mental age or grade level concept became a way to describe a person's developmental level, which was based on specific criterions of development, as compared to the developmental level achieved by the majority of his peers. Normative testing is most simply described as the process of defining an individual's level of skill attainment and relating that attained level to a pre-established group standard. Normative testing data is interpreted through converting raw scores to standard score equivalents of percentiles, grade equivalent, scaled scores, stanines, etc., and not by the degree of consistency nor hierarchial attainment of skill of the students. Gross normative test data results are derived through sampling the student's ability to answer questions representative of the total range of skills inherent within the specified outcome desired. Criterion-referenced testing, on the other hand, is designed to determine the student's ability to perform a specified skill within a comprehensive task. This involves the ability to systematically analyze any given task as to the component parts necessary to perform successfully the determined task. It further makes necessary the successful completion of each hierarchial component before continuing to the next higher developmental stage. Therefore, measurement is done by assessing the child's ability to consistently perform at a given level of quality.

In every unit that a teacher wishes to teach specific objectives are designed as significant parts of the unit. To measure the

student's degree of development or growth, a teacher-made test is used. This test is comprised of specific items that measures behaviors deemed essential to the understanding of the unit. From this test, a teacher receives insights as to the level of knowledge a student has gained as well as insights for the student's success potential in the next higher level of development based upon his current knowledge of necessary sub-skills within the teaching hierarchial pardigm. The teacher-made test is then used as criterion-referenced in the measuring of a student's degree of achievement of specified behaviors within a delimited parameter. This teacher-made test becomes normative based when the teacher assigns a numerical quantity to each child's paper for the purpose of grade distribution. When qualitative distributions are assigned from a criterion-referenced test the purpose served becomes of as much value as the normative-based test. Thus, we as teachers must make a value judgment as to what degree or level the child has achieved as well as his proficiency of the material learned. By placing a proficiency rate, e.g. 70%, on a given skill or sub-skill, we are comparing the child to a normative group proficiency and not to his ability to perform on the task. For example, related sub-skills can be taught as a building process without the attained proficiency rate. If sub-skills are delineated, each lower sub-skill will increase in proficiency as each higher sub-skill is introduced. By using a criterion-referenced test an assessment of a child's growth can be measured and evaluated while attempting the next hierarchial

level of functioning within a specified learning process.

Criterion-referenced tests should be thought of in terms of the educational system; if any skill or task is worth teaching it is worth measuring for growth. Measuring the growth can be accomplished through either positive or negative behavioral observations. Positive behavioral observations can be accomplished by establishing behavioral objectives that are measurable by successfully meeting each criterion in the objectives. Success is measured upon the consistency of growth. Negative observations are accomplished by measuring the lessening of error frequency within each of the objectives.

By establishing behavioral objectives, two prominent teaching tasks are excellent for criterion-referencing. These tasks are teacher-made tests and classwork. Classwork may need to be developed for assessing the child's educational strengths and weaknesses. This may need to be broken down not only in sequential learning steps but also those skills that are needed to do the task. If the behavior a child is displaying can be delineated to minute points, criterion-referenced testing can be of tremendous value.

For the skilled psychometrician, normative based testing is a limited sampling of behavior that is done according to specific prerequisites. The use of norm testing is focused upon that of sorting and selection of students for placement into the educational

curricula that have the potential for meeting the child's educational needs. Meaningful information must be passed between the psychometrician and teacher. Terminology such as "Similarities", "Block Design", "Visual Sequential Memory", and "Auditory Closure" have little meaning to those working with the child if this behavior is not noticeable or defined so that all have a common reference point in viewing the child. When the child is compared to the norming population and to the norm of the test, he is forced into a comparative situation which gives little indication of specific skill deficiency. Thus, when the child is compared to the norm we are doing very little to determine where his strengths and limitations are within the learning process and what remedial activities should be undertaken. By doing a detailed analysis of errors made on normative testing, the examiner can get a much more detailed outlook of the child than what the actual standard scores relate. With this approach, the skilled examiner is better able to relate testing observation to behavior displayed within the classroom setting as well as viewing each test or subtest as to its component parts of incoming stimulus, understanding stimuli or expressive reaction. Thus, the skilled examiner can use a criterion-referenced approach for determination of specific objectives lacking in development. Through observation of the child the skilled teacher must also view a given task in the form of task analysis. From this, both the examiner and teacher can determine the instructional

level for the child and both will be in agreement as to the remedial approach used to best profit the child.

By using a combination of normative based testing and criterion-referenced testing, determination of the child's needs can be focused upon. This can be viewed as a check and balance situation. If for example it is found that a child displaying a particular behavior in formal testing, it serves little remedial purpose if it is not observable within the classroom or affecting the child's total classroom performance. Thus, if synthesizing is not affected, the diminutive behavior may not need to be worked with. Through criterion-referenced testing, evaluation of the synthesizing can be assessed. If it is found that a child has difficulty interpreting the correct word meaning within the sequence of directions, criterion-referenced testing can show this as transferring into the classroom with the development of a checklist that is comprised of behavior sampled in normative testing situation.

These two forms of testing need not be used against each other. They are a means of bringing the teacher and psychometrist closer in their dealings with a child's need. When used together the determination of activities to maximize student learning can be made clearer. Neither can nor should stand on their own; but, both should be used to compliment the other.